



US006298167B1

(12) **United States Patent**
Martucci et al.

(10) **Patent No.:** **US 6,298,167 B1**
 (45) **Date of Patent:** **Oct. 2, 2001**

(54) **VECTOR WAVELET ORGANIZATION**
 (75) Inventors: **Stephen A. Martucci**, Grass Valley, CA (US); **Iraj Sodagar**, Princeton; **Ya-Qin Zhang**, Cranbury, both of NJ (US)

(73) Assignees: **Sarnoff Corporation**, Princeton, NJ (US); **Sharp Corporation**, Osaka (JP)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/177,219**

(22) Filed: **Oct. 22, 1998**

Related U.S. Application Data

- (62) Division of application No. 08/736,114, filed on Oct. 24, 1996.
 (60) Provisional application No. 60/020,852, filed on Jun. 28, 1996, and provisional application No. 60/007,012, filed on Oct. 25, 1995.
 (51) Int. Cl.⁷ **G06K 9/36**
 (52) U.S. Cl. **382/248; 382/240; 375/240.19; 375/240.21**
 (58) Field of Search **382/251, 252, 382/253, 166, 240, 248, 249, 250; 375/240.24, 240.19, 240.21**

(56) References Cited

U.S. PATENT DOCUMENTS

- 5,321,776 * 6/1994 Shapiro 382/251
 5,384,725 1/1995 Coifman et al. 364/807
 5,396,237 * 3/1995 Ohta 341/50
 5,448,297 9/1995 Alattar et al. 348/415
 5,491,513 2/1996 Wickstrom et al. 348/390

- 5,509,089 4/1996 Ghoshal 382/236
 5,526,299 6/1996 Coifman et al. 364/807
 5,546,477 8/1996 Knowles et al. 382/242
 5,563,960 10/1996 Shapiro 382/239
 5,982,938 * 11/1999 Dube 382/240

OTHER PUBLICATIONS

Cosman et al. "Tree-structured vector quantization with significance map for wavelet image coding" Data Compression Conference, Mar. 1995. Dcc '95 Proc. pp. 33-41.*
 Banham et al. "A wavelet transform image coding technique with a quadtree structure" IEEE International Conference on Acoustics, Speech, and Signal Processing, Mar. 1992. vol. 4, pp. 653-656.*

* cited by examiner

Primary Examiner—Amelia M. Au

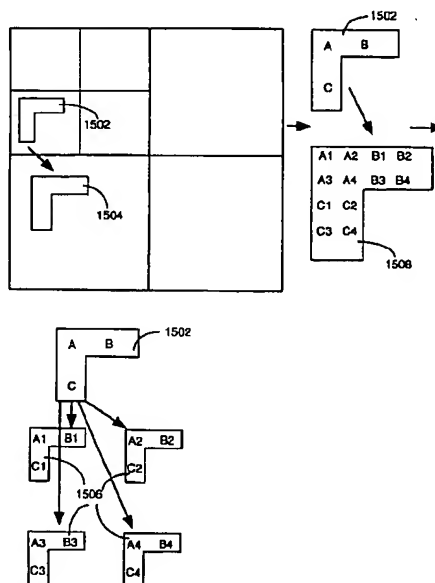
Assistant Examiner—Jingge Wu

(74) Attorney, Agent, or Firm—William J. Burke

(57) ABSTRACT

Apparatus and method for encoding zerotrees in a wavelet-based coding technique. The method uses a depth-first pattern for traversing the zerotree, i.e., each branch of the tree, from parent to child to grandchild and so on, is fully traversed before a next branch is traversed. The depth-first tree traversal pattern is used to quantize the coefficients of the tree as well as to assign symbols to the quantized coefficients. The method assigns one of three symbols to each node: ZEROTREE ROOT, VALUED ZEROTREE ROOT, and VALUE. By using three symbols and the efficient tree traversal pattern, the method is substantially more efficient at encoding a zerotree than the prior art. Additionally, this concept is applied to the encoding of "vector" zerotrees.

13 Claims, 12 Drawing Sheets



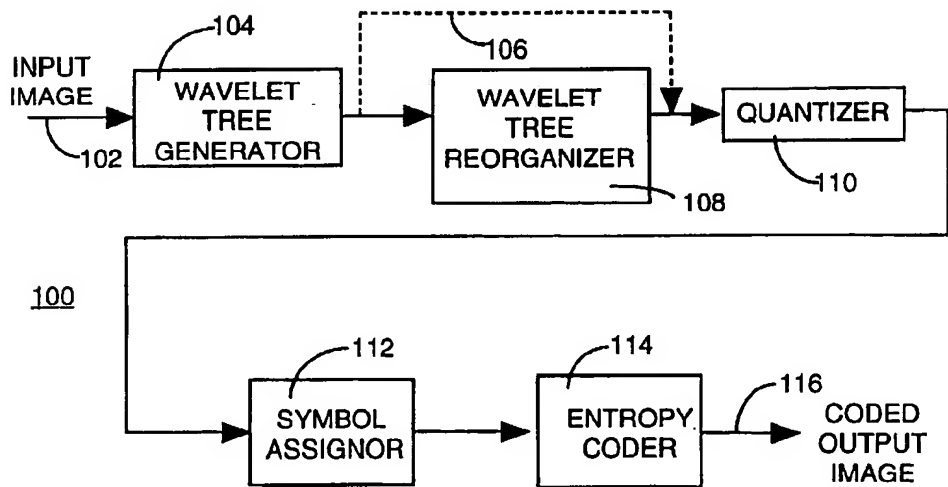


FIG. 1

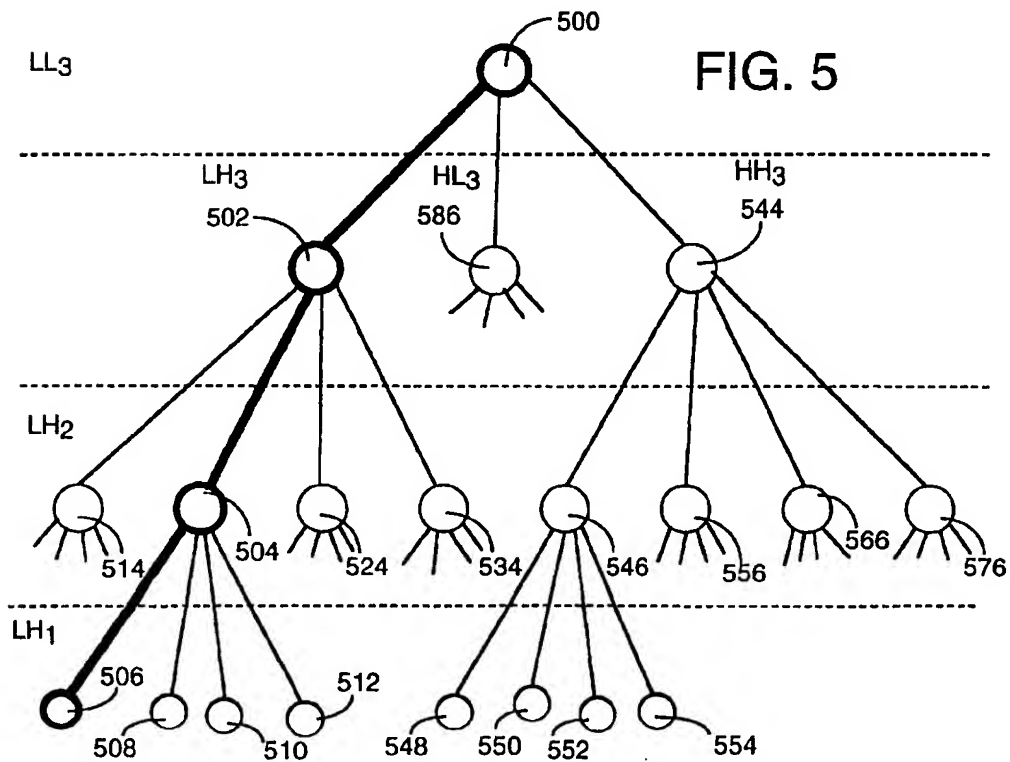


FIG. 5

FIG. 2

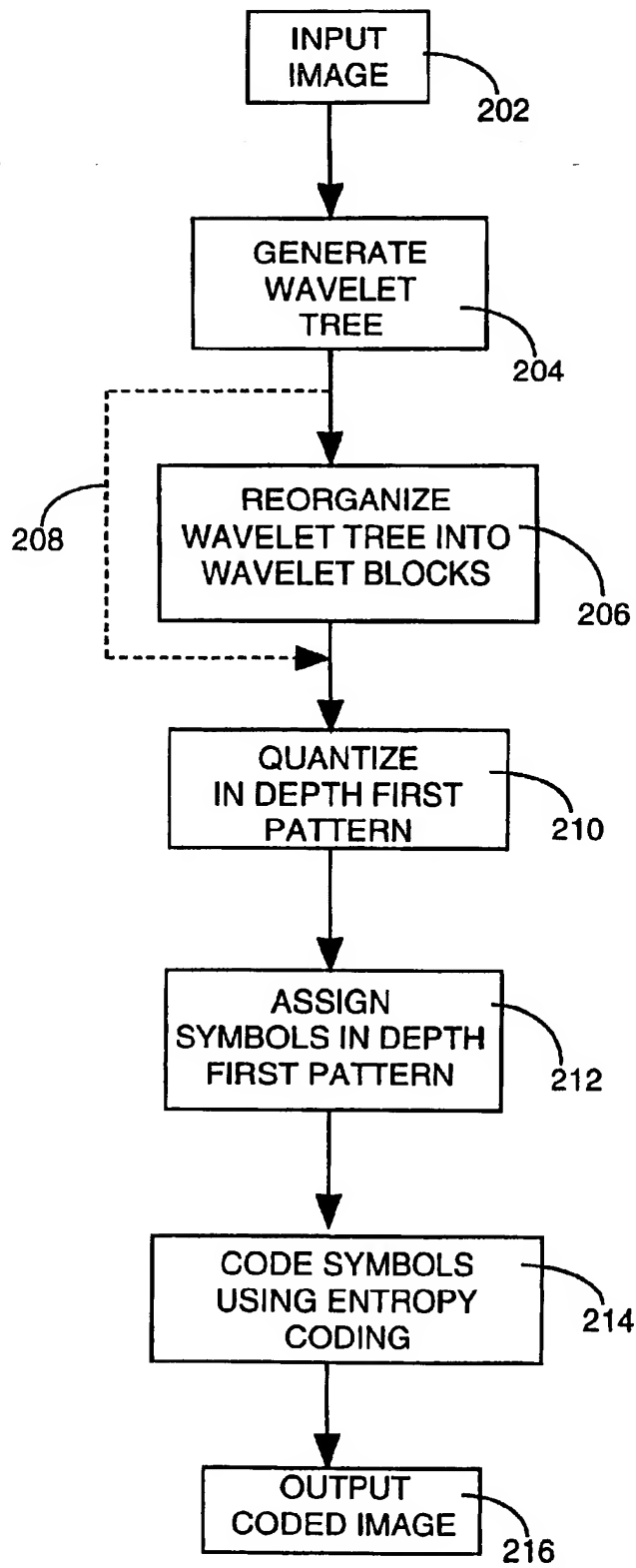


FIG. 3

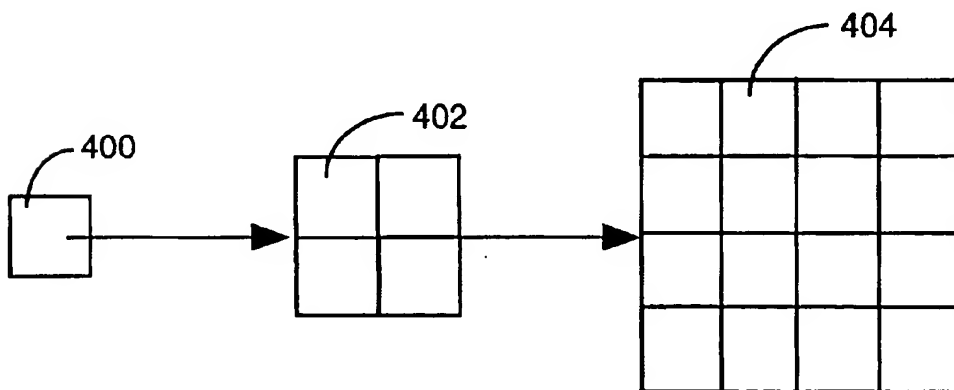
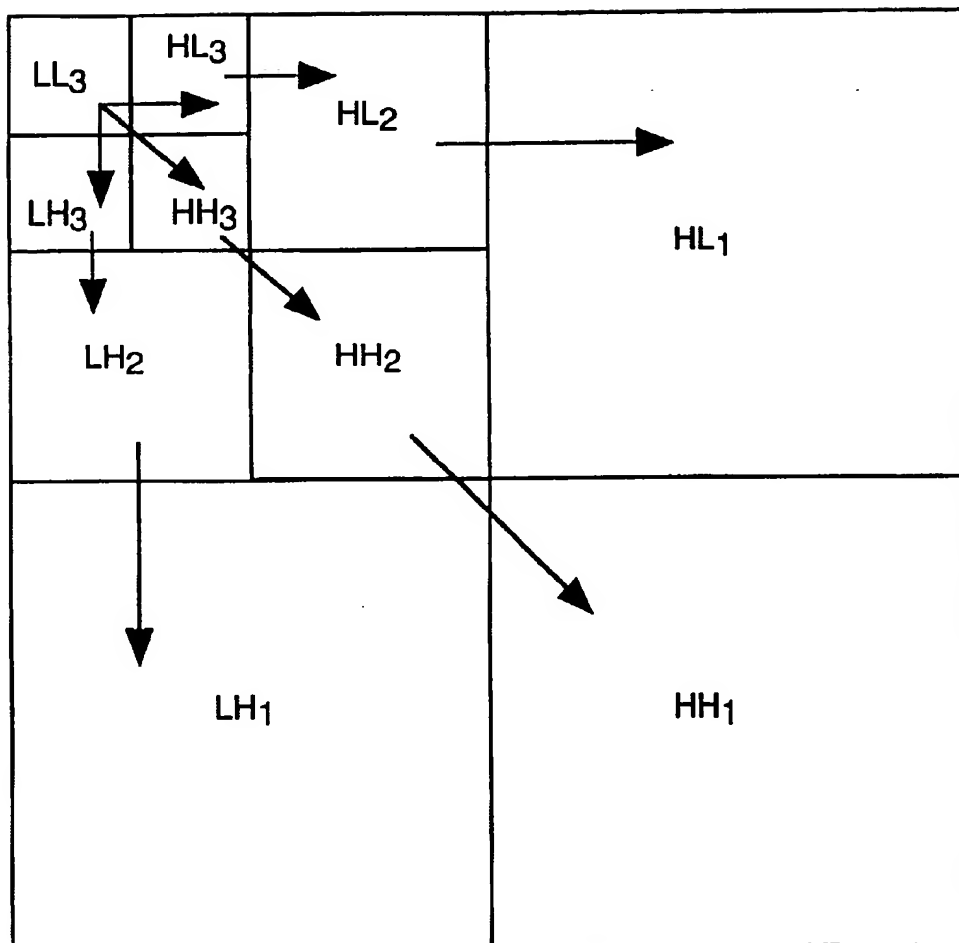
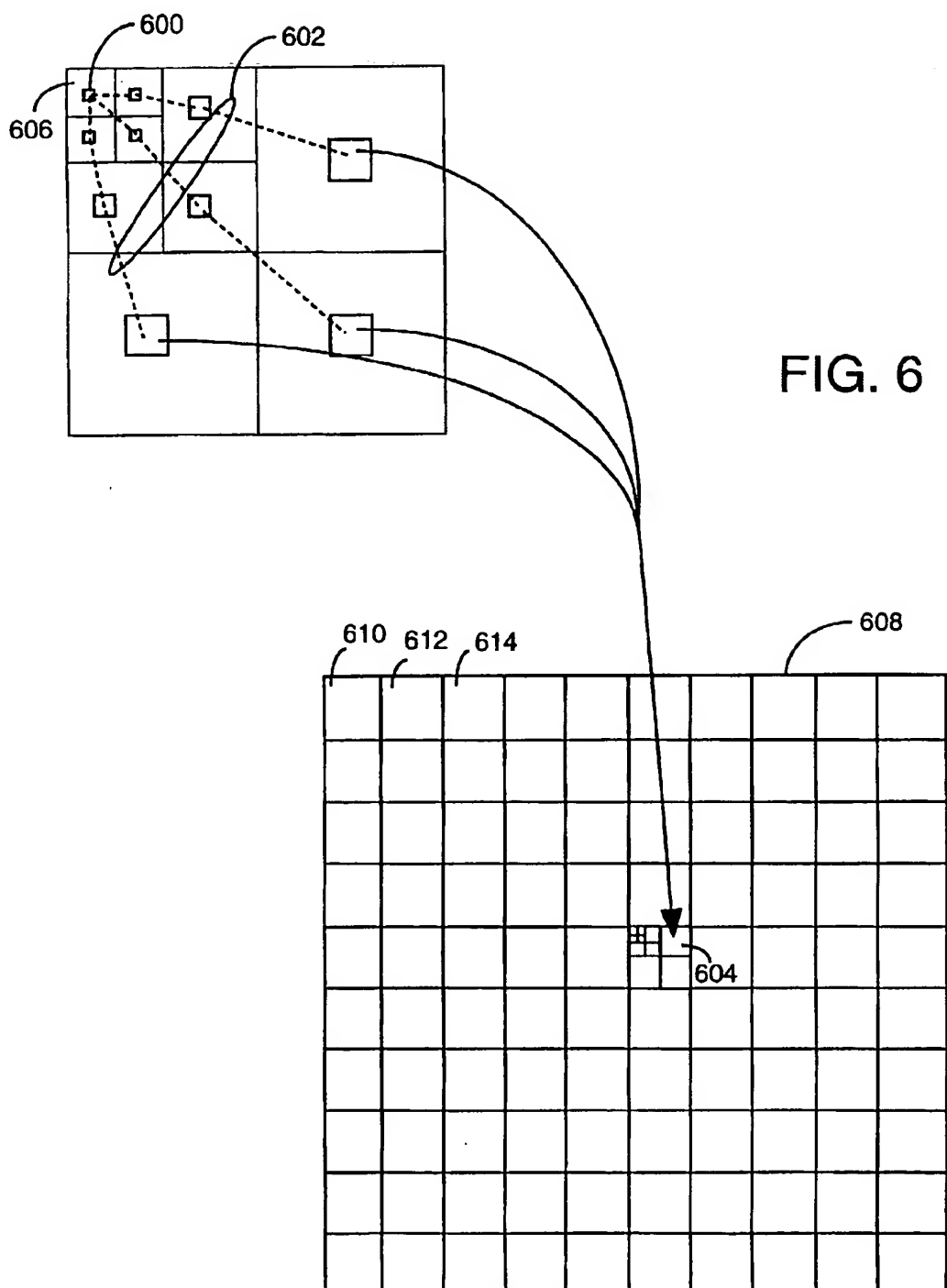


FIG. 4



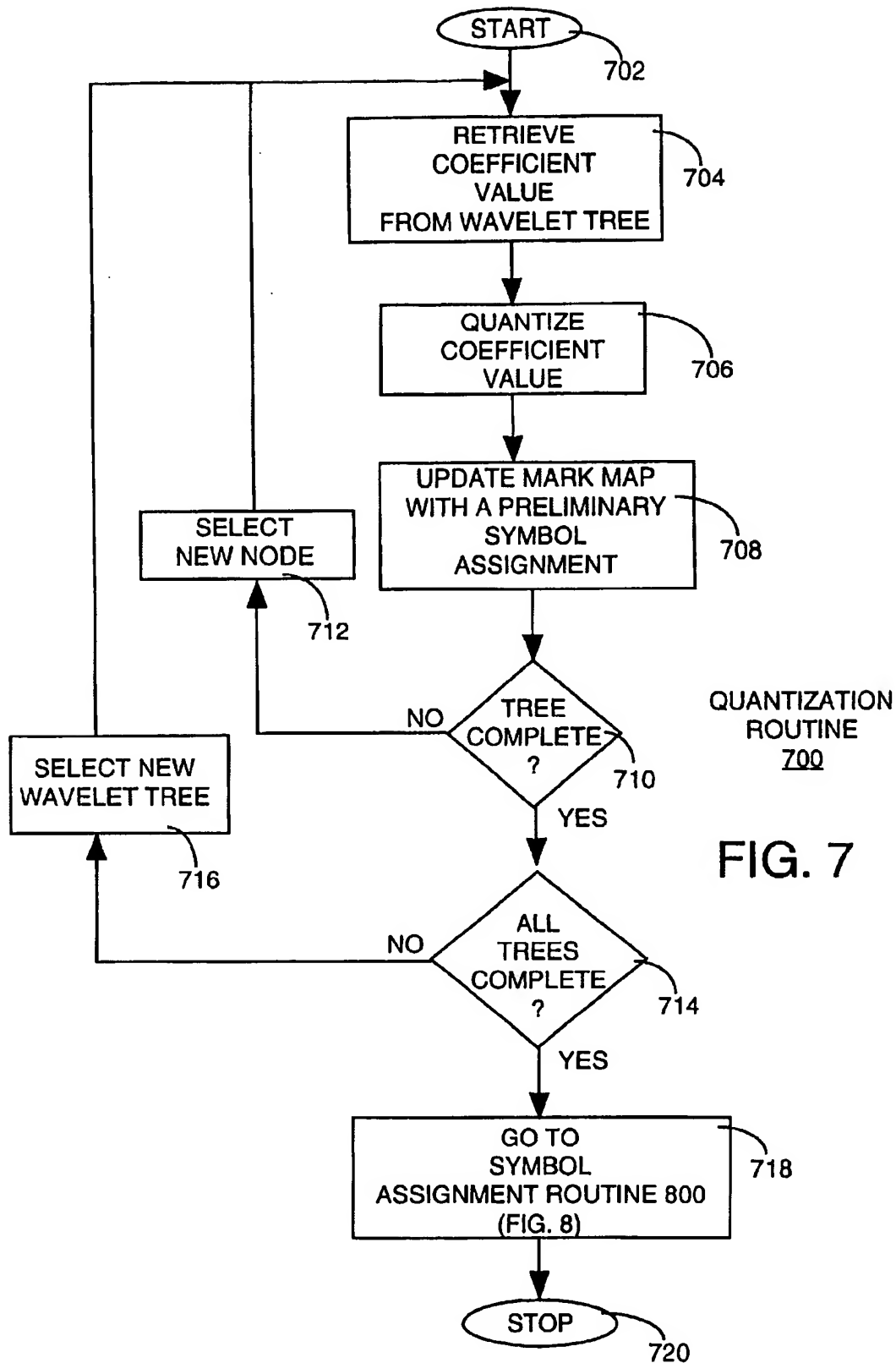


FIG. 8

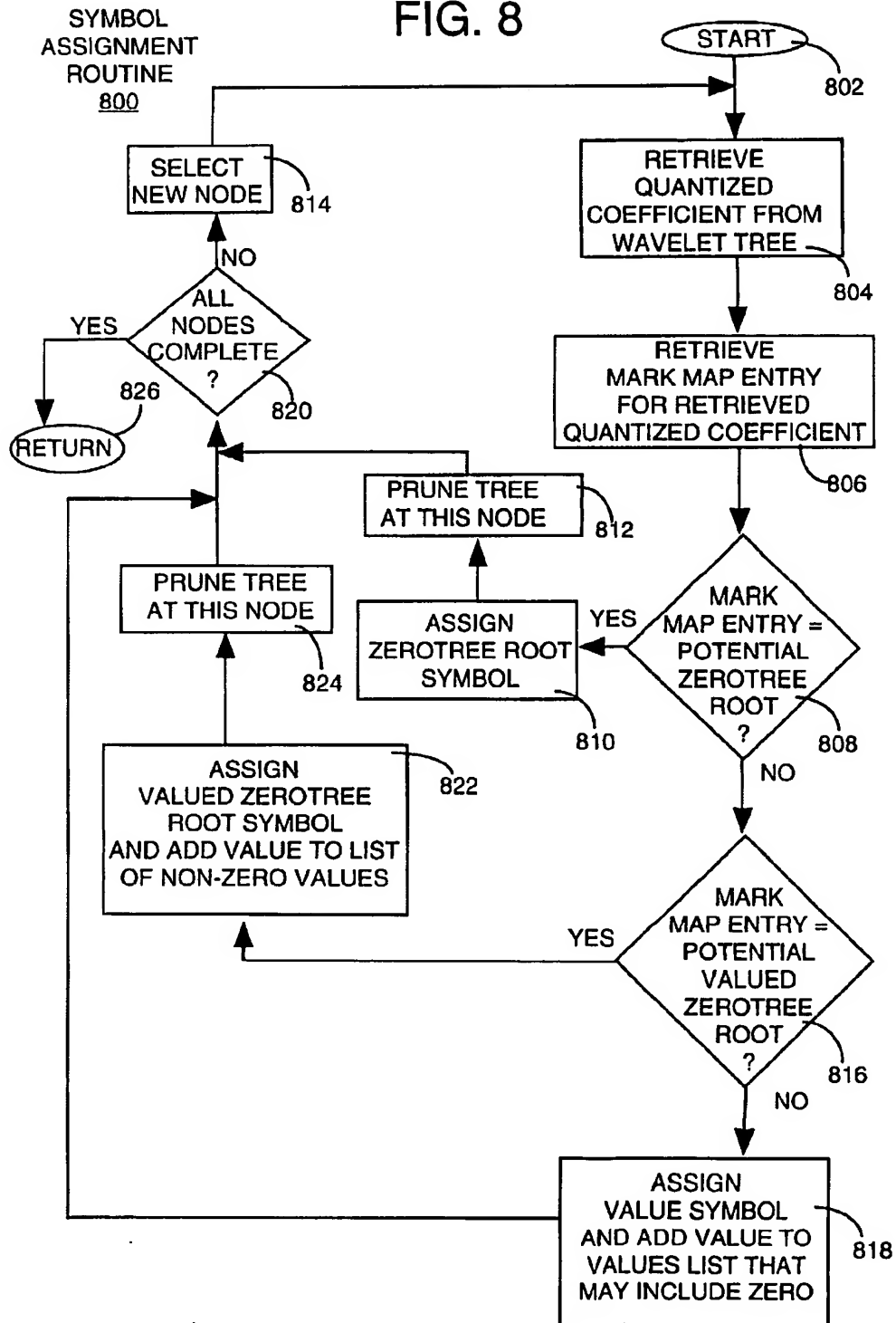


FIG. 9

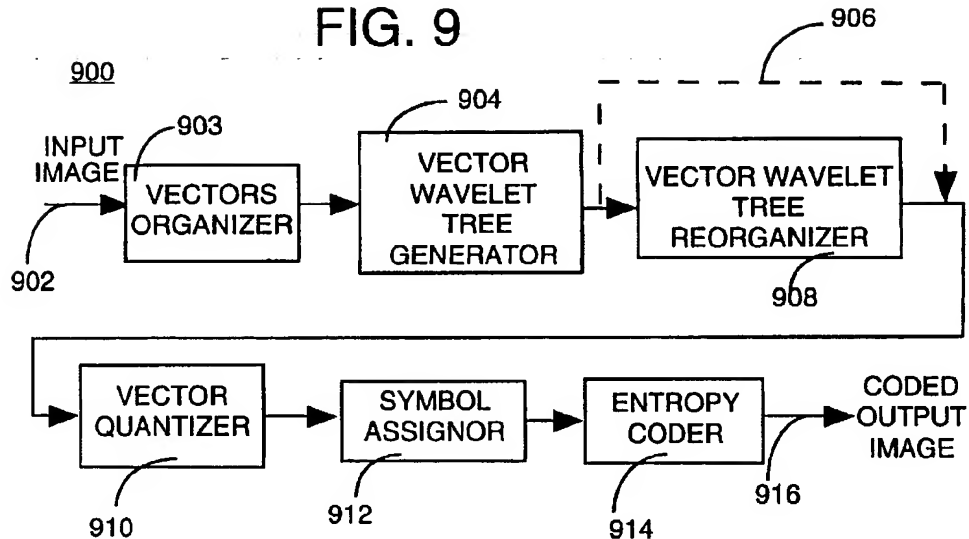


FIG. 10

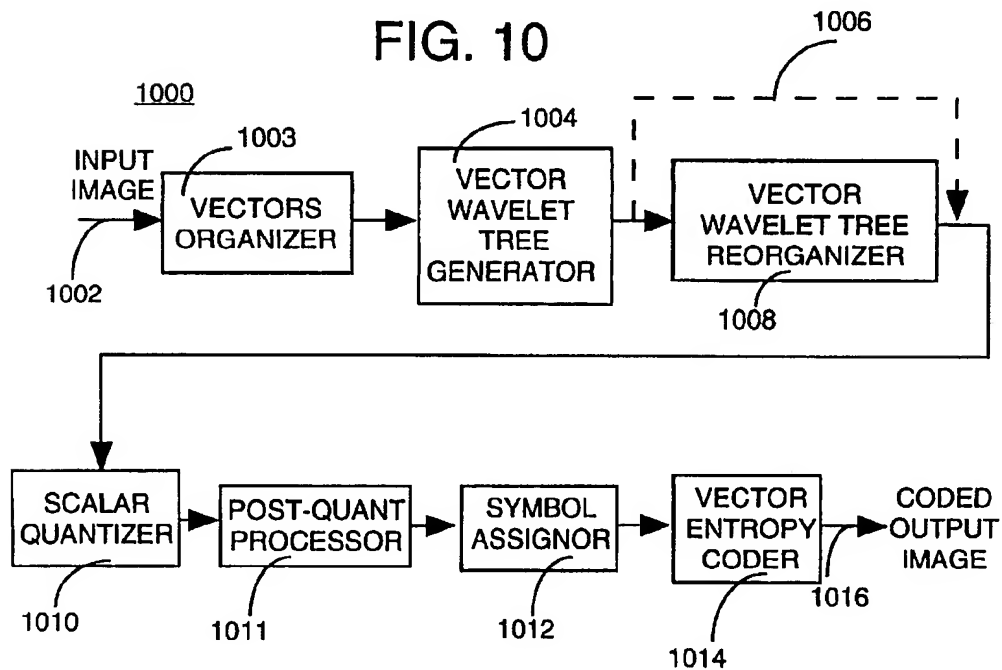


FIG. 11

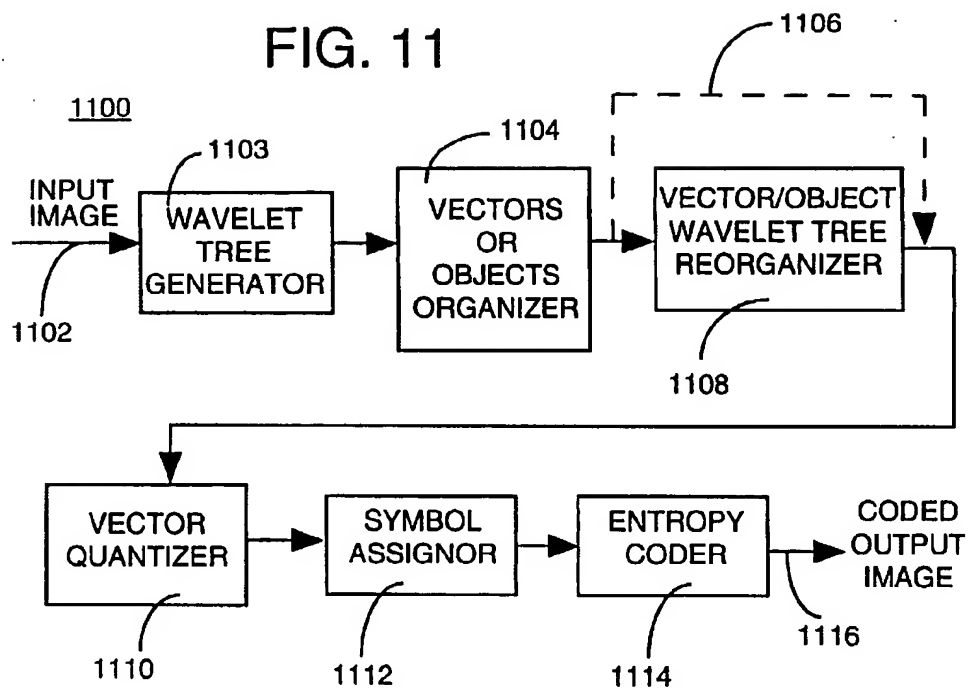
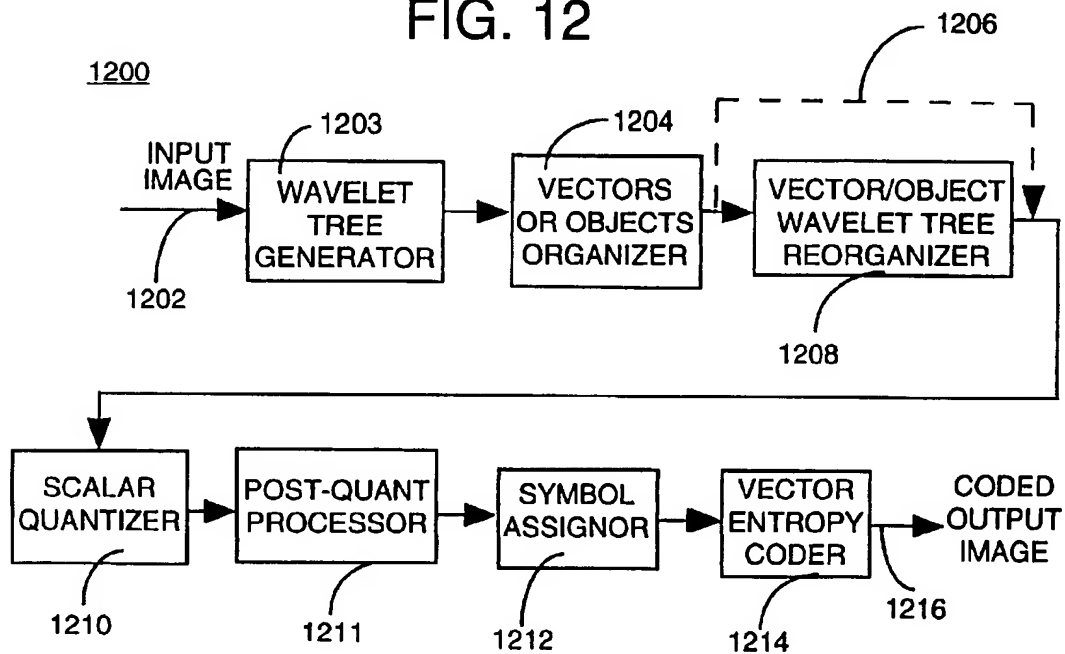


FIG. 12



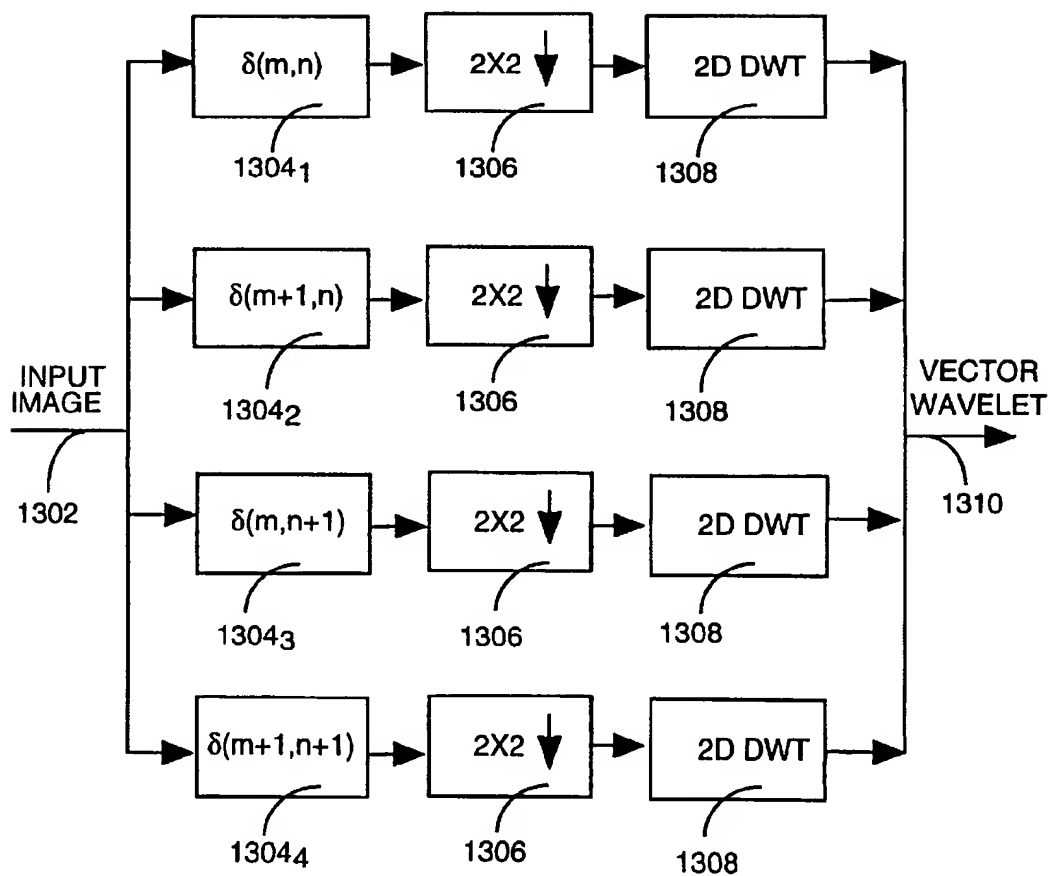


FIG. 13

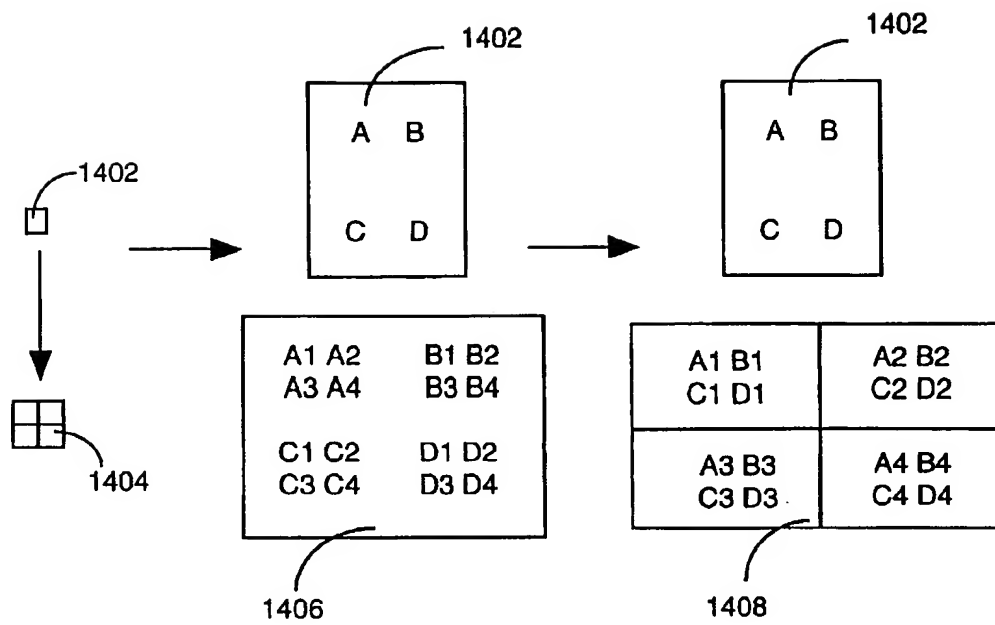


FIG. 14

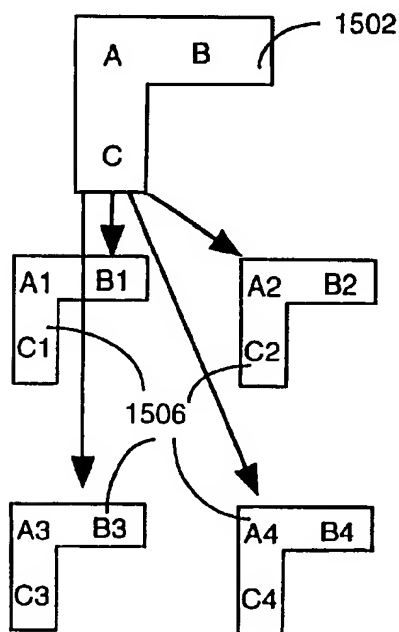
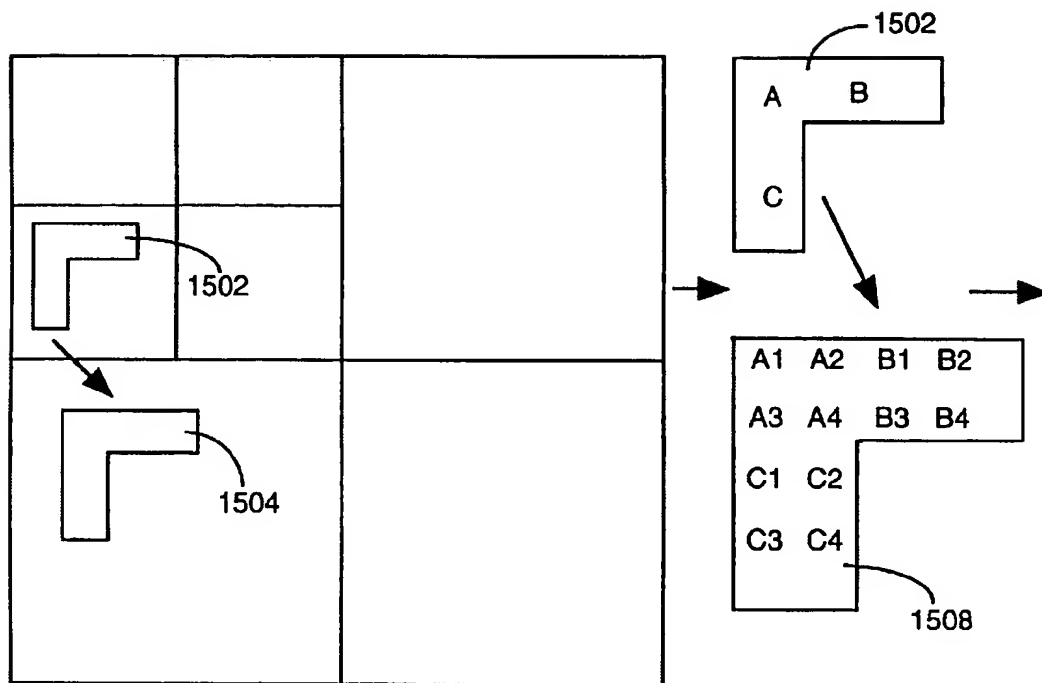
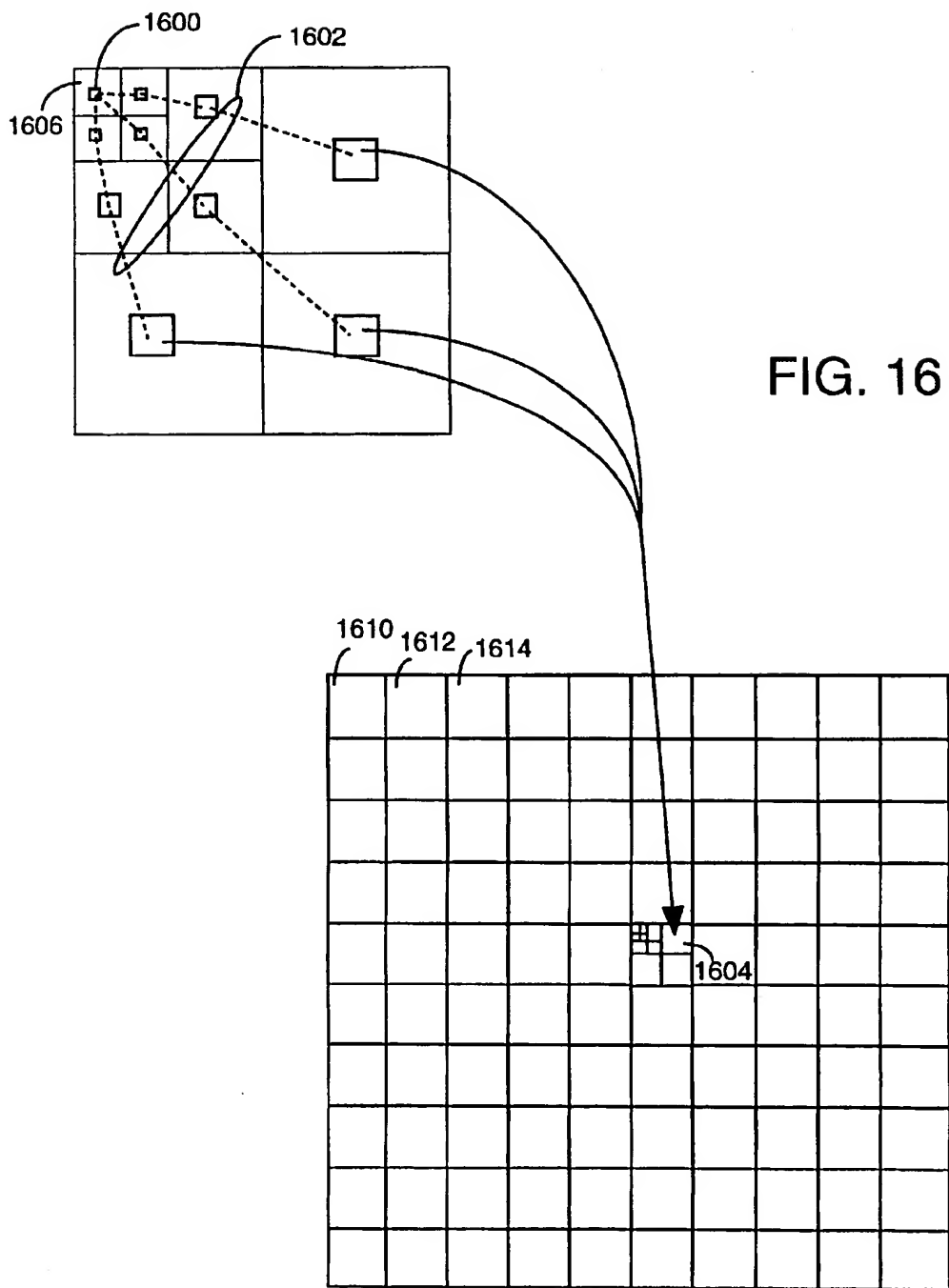


FIG. 15



VECTOR WAVELET ORGANIZATION

This application claims the benefit of U.S. Provisional Application No. 60/007,012 filed Oct. 25, 1995.

This application is a divisional of application Ser. No. 08/736,114, filed Oct. 24, 1996, currently pending, which claims the benefit of the filing date of provisional application No. 60/020,852, filed Jun. 28, 1996.

The invention relates to image processing systems and, more particularly, to image processing systems that utilize wavelet-based coding techniques.

BACKGROUND OF THE DISCLOSURE

Data compression systems are useful for representing information as accurately as possible with a minimum number of bits and thus minimizing the amount of data which must be stored or transmitted in an information storage or transmission system. One of the primary means of compression used in the art removes redundant information from the original data without significantly impacting the quality of the decompressed data when compared to the original data.

One such compression technique appears in the Proceedings of the International Conference on Acoustics, Speech and Signal Processing, San Francisco, Calif. March 1992, volume IV, pages 657-660, where there is disclosed a signal compression system which applies a hierarchical subband decomposition, or wavelet transform, followed by the hierarchical successive approximation entropy-coded quantizer incorporating zerotrees. The representation of signal data using a multiresolution hierarchical subband representation was disclosed by Burt et al. in IEEE Trans. on Commun., Vol Com-31, No. 4, April 1983, page 533. A wavelet pyramid, also known as critically sampled quadrature-mirror filter (QMF) subband representation, is a specific type of multiresolution hierarchical subband representation of an image. A wavelet pyramid was disclosed by Pentland et al. in Proc. Data Compression Conference Apr. 8-11, 1991, Snowbird, Utah. A QMF subband pyramid has been described in "Subband Image Coding", J. W. Woods ed., Kluwer Academic Publishers, 1991 and I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics (SIAM): Philadelphia, Pa., 1992.

Wavelet transforms, otherwise known as hierarchical subband decomposition, have recently been used for low bit rate image compression because such decomposition leads to a hierarchical multi-scale representation of the source image. Wavelet transforms are applied to an important aspect of low bit rate image coding: the coding of a binary map (a wavelet tree) indicating the locations of the non-zero values, otherwise known as the significance map of the transform coefficients. Using scalar quantization followed by entropy coding, in order to achieve very low bit rates, i.e., less than 1 bit/pel, the probability of the most likely symbol after quantization—the zero symbol—must be extremely high. Typically, a large fraction of the bit budget must be spent on encoding the significance map. It follows that a significant improvement in encoding the significance map translates into a significant improvement in the compression of information preparatory to storage or transmission.

To accomplish this task, a new structure called a zerotree has been developed. A wavelet coefficient is said to be insignificant with respect to a given threshold T , if the coefficient has a magnitude less than or equal to T . The zerotree is based on the hypothesis that if a wavelet coefficient at a coarse scale is insignificant with respect to a

given threshold T , then all wavelet coefficients of the same orientation in the same spatial location at finer scales are likely to be insignificant with respect to T . Empirical evidence suggests that this hypothesis is often true.

More specifically, in a hierarchical subband system, with the exception of the highest frequency subbands, every coefficient at a given scale can be related to a set of coefficients at the next finer scale of similar orientation according to a structure called a wavelet tree. The coefficients at the coarsest scale will be called the parent nodes, and all coefficients corresponding to the same spatial or temporal location at the next finer scale of similar orientation will be called child nodes. For a given parent node, the set of all coefficients at all finer scales of similar orientation corresponding to the same location are called descendants. Similarly, for a given child node, the set of coefficients at all coarser scales of similar orientation corresponding to the same location are called ancestors. With the exception of the lowest frequency subband, all parent nodes have four child nodes. For the lowest frequency subband, the parent-child relationship is defined such that each parent node has three child nodes.

Nodes are scanned in the order of the scales of the decomposition, from coarsest level to finest. This means that no child node is scanned until after its parent and all other parents in all subbands at the same scale as that parent have been scanned. This is a type of modified breadth-first, subband by subband, traversal performed across all the wavelet trees defined by the coefficients of the wavelet transform of the two-dimensional data set.

Given a threshold level to determine whether or not a coefficient is significant, a node is said to be a ZEROTREE ROOT if 1) the coefficient at a node has an insignificant magnitude, 2) the node is not the descendant of a root, i.e., it is not completely predictable from a coarser scale, and 3) all of its descendants are insignificant. A ZEROTREE ROOT is encoded with a special symbol indicating that the insignificance of the coefficients at finer scales is completely predictable. To efficiently encode the binary significance map, four symbols are entropy coded: ZEROTREE ROOT, ISOLATED ZERO, and two non-zero symbols, POSITIVE SIGNIFICANT and NEGATIVE SIGNIFICANT.

U.S. Pat. No. 5,412,741 issued May 2, 1995 and herein incorporated by reference discloses an apparatus and method for encoding information with a high degree of compression. The apparatus uses zerotree coding of wavelet coefficients in a much more efficient manner than any previous techniques. The key to this apparatus is the dynamic generation of the list of coefficient indices to be scanned, whereby the dynamically generated list only contains coefficient indices for which a symbol must be encoded. This is a dramatic improvement over the prior art in which a static list of coefficient indices is used and each coefficient must be individually checked to see whether a) a symbol must be encoded, or b) it is completely predictable.

The apparatus disclosed in the '741 patent uses a method for encoding information comprising the steps of forming a wavelet transform of the image, forming a zerotree map of the wavelet coefficients, encoding the significant coefficients on an initial dominant list from the coarsest level of the transform and the children of those coefficients whose indices are appended to the dominant list as the coefficient of the parent is found to be significant, reducing the threshold, refining the estimate of the value of the significant coefficients to increase the accuracy of the coded coefficients, and cycling back to scan the dominant list anew at the new, reduced threshold.

To accomplish the iterative process, the method of the '741 patent is accomplished by scanning the wavelet tree subband by subband, i.e., all parent nodes are coded, then all children, then all grandchildren and so on and they are encoded bit-plane by bit-plane. As the process iterates through the wavelet tree representation of the image, this apparatus codes one of four symbols within the zerotree map. Any improvement in the speed at which a wavelet tree is processed would be advantageous.

Therefore, there is a need in the art for an improved method of classifying and coding the nodes of a wavelet tree that leads to more efficient coding and rapid processing.

SUMMARY OF THE INVENTION

The present invention is apparatus and a concomitant method of encoding zerotrees in a wavelet-based coding technique. Specifically, the invention uses a depth-first pattern for traversing the zerotree, i.e., each branch of the tree, from parent to child to grandchild and so on, is fully traversed before a next branch is traversed. The depth-first tree traversal pattern is used to quantize the coefficients of the tree as well as to assign symbols to the quantized coefficients. Additionally, the invention assigns one of three symbols to each node: ZEROTREE ROOT, VALUED ZEROTREE ROOT, and VALUE. By using three symbols and the efficient tree traversal pattern, the invention is substantially more efficient at encoding a zerotree than the prior art. Additionally, this concept is also applied to the encoding of "vector" zerotrees.

BRIEF DESCRIPTION OF THE DRAWINGS

The teachings of the present invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

FIG. 1 is a block diagram of an image encoder of the present invention;

FIG. 2 is a flowchart illustrating the encoding method of the encoder shown in FIG. 1;

FIG. 3 is a schematic illustration of parent-child dependencies of subbands in an image decomposed to three scales within a wavelet tree;

FIG. 4 depicts the parent-child relationship for three generations of a subsampled image;

FIG. 5 depicts a schematic representation of the interrelation of various nodes within a wavelet tree;

FIG. 6 depicts a wavelet block representation of a wavelet tree;

FIG. 7 depicts a flowchart of a quantization routine executed by the present invention;

FIG. 8 depicts a flowchart of a symbol assignment routine executed by the present invention;

FIG. 9 depicts a block diagram of one embodiment of a vector-wavelet transform;

FIG. 10 depicts a block diagram of another embodiment of a vector-wavelet transform;

FIG. 11 depicts a block diagram of one embodiment of a scalar-wavelet transform followed by vectors or object grouping;

FIG. 12 depicts a block diagram of another embodiment of a scalar-wavelet transform followed by vectors or object grouping;

FIG. 13 depicts an example of a vector wavelet;

FIG. 14 depicts a vector parent-child relationship;

FIG. 15 depicts an object parent-child relationship; and

FIG. 16 depicts the reorganization of the coefficient vectors of each wavelet tree to form a "wavelet block".

To facilitate understanding, identical reference numerals have been used, where possible, to designate identical elements that are common to the figures.

DETAILED DESCRIPTION

FIG. 1 depicts a block diagram of an encoder 100 of the present invention and FIG. 2 depicts a flowchart representation of the operation of the encoder 100 of FIG. 1. To best understand the invention, the reader should simultaneously consult both FIGS. 1 and 2 while reading the following description of the invention.

The encoder 100 contains a wavelet tree generator 104, an optional wavelet tree reorganizer 108, a quantizer 110, a symbol assignor 112, and an entropy encoder 114. Each of these components is connected in series to process an image at port 102 into a coded output image at port 116. The input image is typically a pixelated (digitized) photographic image as can be produced from an image scanner or a computer graphics system. However, the input image can also be a frame within a series of frames of video images or a motion compensated residuals frame produced by a video encoding system. In general, the invention processes any form of digitized image or portion thereof. Thus, the method of operation generally begins at step 202 with the input of an "image", i.e., any form of two-dimensional data.

The wavelet tree generator 104 performs (at step 204) a wavelet hierarchical subband decomposition to produce a conventional wavelet tree representation of the input image. To accomplish such image decomposition, the image is decomposed using times two subsampling in each of two-dimensions into high horizontal-high vertical (HH), high horizontal-low vertical (HL), low horizontal-high vertical (LH), and low horizontal-low vertical (LL), frequency subbands. The LL subband is then further subsampled times two in each of two dimensions to produce a set of HH, HL, LH and LL subbands. This subsampling is accomplished recursively to produce an array of subbands such as that illustrated in FIG. 3 where three subsamplings have been used. Preferably four or more subsamplings are used in practice. The parent-child dependencies between subbands are illustrated as arrows pointing from the subband of the parent nodes to the subbands of the child nodes. The lowest frequency subband is the top left LL₃, and the highest frequency subband is at the bottom right HH₁. In this example, all child nodes have one parent. A detailed discussion of subband decomposition is presented in J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", IEEE Trans. on Signal Processing, Vol. 41, No. 12, pp. 3445-62, December 1993.

FIG. 4 depicts the parent-child relationship for three generations of a subsampled image. A single parent node 400 has four child nodes 402 corresponding to the same region in the image with times four subsampling, i.e., times two subsampling in each of two dimensions. Each child node 402 has four corresponding next generation child nodes 404 with a further times four subsampling. The relationship, or data structure, that relates a parent node to its children and grandchildren is a wavelet tree. Note that each pel or pixel in the low-low subband has a tree associated with it. However, the plurality of trees that extend from the low-low subband taken together are generally discussed in the art as "the" wavelet tree for the image. This disclosure will also follow this nomenclature.

5

As shown in FIGS. 1 and 2, the quantizer 110 quantizes (at step 210) the coefficients of the wavelet tree in a "depth-first" pattern. A depth-first pattern is very dissimilar from the prior art encoding systems in that the prior art encoding systems traverse the wavelet tree in a modified breadth-first pattern across all subbands, i.e., quantize all parent nodes, then all children, then all grandchildren and so on. In contrast, the inventive method traverses each tree in "depth-first" order, i.e., from the root in the low-low subband (LL_3) through the children. FIG. 5 depicts the depth-first pattern used to traverse each tree. For example, beginning at node 500 in LL_3 and following the bold path, the inventive depth-first process proceeds to node 502 in subband LH_3 and then to node 504 in subband LH_2 . From node 504, the depth-first traversal process successively continues to nodes 506, 508, 510 and 512 within subband LH_1 , i.e., all the children of node 504, then continues on to the siblings of 504 (514, 524, 534) where the four children of each sibling are traversed before the next sibling and its children. Once this entire branch of the tree is traversed, the traversal process proceeds to another child node of node 500, for example, node 544. From that node, the depth-first traversal process proceeds to nodes 546, 548, 550, 552 and 554 before going on to node 556 and so on.

As each branch is traversed, the coefficients are quantized into discrete values. Any quantization approach can be used with the present invention. The quantization process maps a continuous coefficient value to a discrete value having either a positive value, a negative value or zero value. In sum, in a depth-first scan pattern, children 506, 508, 510, and 512 are scanned after their parent 504 and before any of the neighboring parents 514, 524 and 534. In this manner, all coefficients that represent a given spatial location are scanned, in ascending frequency order from parent 500 to child 502 to grandchild 504 and so on, before the coefficients of the next adjacent spatial location is scanned.

Although the foregoing description of the depth-first scanning pattern was discussed as a "top down" pattern, a depth-first scanning pattern also includes scanning from the bottom up. As such, the quantization can also be accomplished by starting at a tree's "leaves" (the bottom-most nodes) and proceeding up the tree. Using the example of FIG. 5, in a "bottom up" pattern, nodes 506, 508, 510 and 512 would be quantized first, then node 504, and so on up the tree to 500 last. Once that tree was complete, the quantization process would quantize another tree, and another, and so on until all the nodes in all the trees were quantized. As shall be discussed below, the invention operates more efficiently when using a bottoms up pattern than the top down.

To facilitate this depth-first scanning pattern, the invention reorganizes the quantized coefficients of each wavelet tree to form a "wavelet block". As shown in FIGS. 1 and 2, the reorganization is accomplished (at step 206) in the wavelet tree reorganizer 108 prior to quantization.

FIG. 6 schematically depicts a wavelet block 604 that is generated by the invention. The invention maps a tree 602 extending from a pixel 600 in the low-low band $606 (LL_3)$ in the wavelet tree 602 into a wavelet block 604. Each wavelet block 604 of an image frame 608 comprises those coefficients at all scales and orientations that represent the frame at the spatial location of the block within the frame. The reorganization is accomplished by physically remapping the memory locations of the coefficients to new memory locations that form the wavelet blocks. As such, all the coefficients of a given wavelet block are stored at sequential address locations. Alternatively, the coefficients

6

are not physically rearranged, but are rather remapped into a virtual memory. Thus, an index into the physical memory is created, where the index (virtual memory) has memory locations that are arranged into wavelet blocks. For each access into the index, the address into the index is mapped to a physical memory location where the coefficient is stored. Thus, by a virtual memory approach, the advantages of wavelet blocks are available without physically rearranging the coefficients in memory.

By using a depth-first scanning pattern, each wavelet block is completely scanned to quantize its coefficients before the next block is scanned and so on. For example, block 610 is completely scanned, then block 612, then block 614, and so on in a raster scan pattern through the frame of wavelet blocks. The ordering of blocks does not have to be in a raster scan pattern, but can be any order as desired by the application. This includes object-oriented whereby blocks corresponding to certain objects are scanned and coded before other objects. Since an entire block is located at consecutive memory addresses, the block can easily be scanned in a top down or bottoms up pattern by selecting either the first or last memory entry for a given block and accessing all other addresses in ascending or descending order.

Importantly, with such reorganization, each wavelet block can be assigned a different quantizer scale based on its spatial location in the frame. This permits the quantizer 110 to be allocated specifically for a spatial location of the coefficients and/or in accordance with the frequency band represented by the coefficient. As such, the scale of the quantizer can be different across an image such that the center of the image or certain objects within the image can be more accurately quantized than the edges. Similarly, the quantizer scale could be frequency dependent such that higher frequency (or, for that matter, lower frequencies, middle frequencies, various frequency bands, and the like) can be quantized using a scale that is different from other frequencies.

In a particular application, such as a video-phone, the foreground object(s), e.g., the head and shoulders of a caller, are elected for quantization with more accuracy than the background objects. Consequently, the important information is accurately coded and transmitted, while background information is less accurately coded. The use of wavelet blocks allows for easy bifurcated coding of foreground and background information. Once a region is selected for accurate coding, the invention can flag the blocks that require accurate coding. As such, the quantizer quantizes those blocks using a finer quantization step than is used on the unflagged blocks.

Although wavelet blocks form an intuitive data structure for implementing the invention, use of wavelet blocks is not necessary to implementation of the inventive encoder 100. As shall be discussed below, the conventional tree structure can be used in conjunction with the improved tree traversal process and the improved coding technique of the present invention. As such, FIGS. 1 and 2 depict the optional nature of the reorganizer as path 106 and path 208 which respectively bypass the reorganizer and its associated function.

After quantization, at each node of the tree, the quantized coefficient has either a zero value or a non-zero value. "Zerotrees" exist wherever the coefficient at a node is zero and all its descendants form zerotrees, i.e., all descendant nodes have a zero value. The quantized coefficients of the wavelet tree are efficiently encoded by again scanning each tree in a depth-first manner. Thus, symbol assignor 112

operates (at step 212) by traversing the tree and assigning particular symbols to each node depending upon the node's quantized value as well as the quantized values of each node's descendants.

Specifically, at each node, the inventive method assigns one of three symbols: ZEROTREE ROOT, VALUED ZEROTREE ROOT, and VALUE. A ZEROTREE ROOT denotes a coefficient that is the root of a zerotree. After the scan in which symbols are assigned, the zerotree does not need to be scanned any further because it is known that all coefficients in the tree have the value zero. A VALUED ZEROTREE ROOT is a node where the coefficient has a non-zero value and all four children are ZEROTREE ROOTS. The coding scan of this tree never progresses below this node. A VALUE symbol identifies a coefficient with a value, either zero or non-zero, but also with some descendant somewhere further along the tree that has a non-zero value.

To most efficiently scan the trees to quantize and assign symbols to the nodes, the quantizer operates in conjunction with the symbol assignor. FIG. 7 depicts a detailed flowchart of a quantization routine 700 used to quantize the coefficients of a zerotree and FIG. 8 depicts a detailed flowchart of a symbol assignment routine 800 for assigning symbol values to represent the quantized coefficient values.

The routine 700 begins at block 702 and proceeds to step 704 where a coefficient value is retrieved from a node in a wavelet tree. As shall be discussed below, the quantization routine scans the wavelet tree in a bottom up, depth-first pattern. Thus, the first address is always in the highest frequency subband and, with each iteration through the routine, the routine proceeds up the tree to lower and lower frequency subbands. As the quantized values are generated, the routine keeps track of the quantized values of the child nodes, i.e., are the children valued or zero. At step 706, the routine quantizes the retrieved coefficient value into a positive value, a negative value, or zero value. At step 708, a mark map is updated with a preliminary symbol value for the node associated with the coefficient value just quantized. The mark map symbol depends upon the value of the child nodes as well as the value of the present node. Note that, because the scan is accomplished bottom up, the mark map is not capable of conclusively indicating whether a node is a ZEROTREE ROOT or not. Consequently, after all the nodes are assigned a preliminary symbol value, the tree is scanned again in a top down pattern to conclusively assign symbol values. The mark map is an index of the wavelet tree nodes which is filled by the quantization routine 700. At each address in the mark map, the routine stores a preliminary symbol: potential VALUE, potential VALUED ZEROTREE ROOT, or potential ZEROTREE ROOT. If the quantized coefficient value has a value, the mark map location for that coefficient is marked with a potential VALUE symbol. If the quantized coefficient value is zero value and all of that nodes children are zero valued, then the mark map location is marked with a potential ZEROTREE ROOT. Lastly, if the quantized value has a value and its children are all zero valued, then the mark map location is marked with a potential VALUED ZEROTREE ROOT.

At step 710, the routine queries whether all the nodes in the tree have been quantized. If the query is negatively answered, the routine proceeds to step 712 where a new (next) node in the tree is selected for quantization. The routine then returns to step 704. If the query at step 710 is affirmatively answered the routine proceeds to step 714. The routine queries at step 714 whether all the trees have been quantized. If the query is negatively answered, the routine

selects, at step 716, a new (next) tree or quantization. If the query at step 714 is affirmatively answered, the routine proceeds to step 718. At this point in routine 700, all the nodes in all the trees have been quantized and assigned a preliminary symbol. At step 718, the routine 700 calls symbol assignment routine 800 of FIG. 8. After the symbols are assigned, routine 700 ends at block 720.

Routine 800 scans the trees in a top down pattern, i.e., root to leaves. However, the routine does not need to access every node because the trees are pruned at each occurrence of a ZEROTREE ROOT or a VALUED ZEROTREE ROOT. Specifically, routine 800 is entered at step 802 and proceeds to step 804. At step 804, the routine retrieves a quantized coefficient from the tree of quantized coefficients. At step 806, the routine retrieves the preliminary symbol in the mark map that corresponds to the retrieved coefficient. The routine queries at step 808 whether the preliminary symbol is a potential ZEROTREE ROOT. If the query is affirmatively answered, the routine assigns, at step 810, the ZEROTREE ROOT symbol to the node. Then, at step 812, the routine prunes the tree, i.e., the routine ignores all nodes below this ZEROTREE ROOT node because, by definition, all the nodes have a zero value.

The routine queries at step 820 whether all nodes have been selected. If the query at step 820 is negatively answered, the routine proceeds along the NO path to step 814. At step 814, the routine selects the next node, after any pruned branches are skipped, in the tree such that a top down, depth-first scan is accomplished.

If the query at step 808 is negatively answered, the routine proceeds along the NO path to step 816. At step 816, the routine queries whether the mark map contains a potential symbol of potential VALUED ZEROTREE ROOT. If the query at step 816 is affirmatively answered, the routine, at step 822, assigns a VALUED ZEROTREE ROOT symbol to the node, puts the value on a list of non-zero values, and prunes the tree at step 824. The routine queries at step 820 whether all nodes have been selected. If the query at step 820 is negatively answered, the routine proceeds to step 814. Then the routine, at step 814, selects the next node for symbol assignment, skipping the pruned branches.

If the query at step 816 is negatively answered, the routine assigns, at step 818, a VALUE symbol to the node, and puts a value on the list of values that include the value zero. The routine queries at step 820 whether all nodes have been selected. If the query at step 820 is negatively answered, the routine proceeds to step 814. Then, the routine selects, at step 814, the next node for symbol assignment.

The assignment routine continues until all nodes have had symbols assigned to them. Thus, if the query at step 820 is affirmatively answered, the routine proceeds to step 826, where the routine 800 ends or returns to routine 700.

The symbols and values are encoded (at step 214) using an entropy coder 114, such as a conventional arithmetic coder. One possible way to accomplish encoding is as follows. The symbols are encoded using a three-symbol alphabet. The list of non-zero values that correspond one-to-one to the VALUED ZEROTREE ROOT symbols is encoded using an alphabet that does not include the value zero. The remaining coefficients, which correspond one-to-one to the VALUE symbols, are encoded using an alphabet that does include the value zero. For any node reached in a scan that is a leaf with no children, neither root symbol could apply. Therefore, some bits can be saved by not encoding any symbol for this node and encoding the coefficient using the alphabet that includes the value zero. An illustrative

encoder using a three-symbol coding alphabet for the symbols and a multi-symbol alphabet for the values would follow that disclosed in Witten et al., "Arithmetic Coding for Data Compression", Comm. of the ACM, Vol. 30, No. 6, pp. 520-540, June 1987. In fact, those skilled in the art will realize that the present invention can be modified by simply encoding only the values (or representations of those values) of the coefficients in accordance with the assigned symbols. Namely, only the values of the coefficients are encoded without having to encode the symbols that indicated the importance of the coefficients.

The encoder 100 generates (at step 216) the coded output image at port 116. Through utilization of the present invention, an image is rapidly and efficiently coded using one of three symbols at each node of a wavelet tree plus bits to encode coefficient values.

A plurality of alternate embodiments of the present invention applies the present coding scheme to vector or object zerotree coding for images or video. Namely, the present coding scheme, using a wavelet transform, is applied to the concept of vector (or object) zerotree coding which entails quantization and zerotree entropy coding of the quantized wavelet coefficients. Each of the vector or object zerotree coding embodiments of the present invention constitutes a vector or object approach to using zerotrees to encode wavelet coefficients for images or video residuals. Associated with each node of each wavelet tree is a list of coefficients (a vector) rather than a single coefficient. Tree nodes are scanned, coefficients quantized, and symbols entropy coded all using a method called Vector ZeroTree Entropy coding.

The various vector or object zerotree coding embodiments fall into two categories, scalar-wavelet transformation followed by vectors (or objects) grouping, or vectors (or objects) grouping followed by vector-wavelet transformation. Namely, an image, a video frame, or a frame of motion-compensated video residuals is first either scalar-wavelet transformed or vector-wavelet transformed. FIGS. 11-12 illustrate the first category where the scalar wavelet is first computed, then wavelet coefficients are grouped into vectors or objects after the transformation. In contrast, FIGS. 9-10 illustrate the second category, where the inputs are grouped as vectors and are passed through a vector-wavelet transform, to produce the outputs as vectors. In both cases, the result can be viewed as a frame of vectors or objects of wavelet coefficients.

Briefly, in two embodiments of the invention (FIGS. 9 and 10), vectors are defined by a fixed tiling of the 2-D matrix of image or video pixels or residuals and the wavelet transform computed is a vector wavelet transform. Then, either vector quantization or scalar quantization is combined with vector zerotree entropy coding to yield a coded output image.

In another four embodiments of the invention (FIGS. 11 and 12), a scalar wavelet transform is computed and then either vectors are defined by a fixed tiling of the 2-D matrix of wavelet coefficients at the coarsest scale or objects of arbitrary size and shape are determined by a segmentation algorithm (i.e., producing two embodiments). In turn, the vectors or objects are either vector quantized or scalar quantized in combination with vector zerotree entropy coding to yield a coded output image (i.e., producing another two embodiments for a total of four embodiments). One form of a vector wavelet is disclosed by W. Li and Y.-Q. Zhang, "Vector-Based Signal Processing and Quantization for Image and Video Compression", *Proceedings of the IEEE*, vol. 83, pp. 317-335, February 1995.

FIGS. 13 illustrates an example of a vector wavelet. A vector wavelet of an image or video residual is a transform whereby adjacent samples of the image or video residual are grouped together to form vectors and then the wavelet transform is applied to those vectors. A vector wavelet transform is computed by the following steps: subsampling the frame to obtain a number of subsampled frames of spatially offset frame-representative signals; transforming each of the subsampled frames by a discrete wavelet transform to obtain a corresponding number of frames of wavelet transform coefficient-representative signals; and forming vector-representative signals from corresponding coefficient-representative signals of the respective frames of coefficient-representative signals.

Namely, the frame (input image) 1302 is decomposed into its 2-D polyphase components. Each polyphase component is a spatially-offset frame of the original frame as illustrated by the four (4) delta functions 1304₁-1304₄. Each spatially-offset frame is then subsampled with times four subsampling, i.e., times two subsampling in each of two dimensions by subsampler 1306 to produce the polyphase component. In turn, a wavelet transform 1308 is applied to each polyphase component, and then vectors are formed by grouping together the transform coefficients, one from each wavelet-transformed component, with the same indices. In this manner, vectors of $m \times n$ pixels are transformed into vectors of $m \times n$ coefficients 1310.

Vector wavelet transform can also be viewed as a number of overlaying frames and each vector or object comprises one coefficient from each frame where those coefficients all share the same index. The number of frames is equal to the dimension of the vectors or size of the objects.

The various embodiments are now discussed in detail with reference to FIGS. 9-12. More specifically, FIGS. 11-12 illustrate four embodiments that apply scalar-wavelet transformation followed by vectors (or objects) grouping. Since the encoders of FIGS. 11 and 12 share many similarities, the various embodiments are discussed simultaneously below through the use of numerical labels.

The encoder 1100 (1200) contains a wavelet tree generator 1103 (1203), a vectors or objects organizer 1104 (1204), an optional vector/object wavelet tree reorganizer 1108 (1208), a vector quantizer 1110 (a scalar quantizer 1210 with a post-quantization processor 1211), a symbol assigner 1112 (1212), and an entropy encoder 1114 (a vector entropy encoder 1214). Each of these components is connected in series to process an image at port 1102 (1202) into a coded output image at port 1116 (1216).

The wavelet tree generator 1103 (1203) performs a wavelet hierarchical subband decomposition to produce a conventional wavelet tree representation of the input image 1102 (1202) as discussed above. In turn, vectors or objects organizer 1104 (1204) defines vectors by a fixed grouping of adjacent coefficients in the low-low band of the wavelet transform. Objects are defined by applying a segmentation algorithm to the wavelet coefficients of the low-low band. The segmentation is assumed to be complete and nonoverlapping. Therefore, each coefficient in the low-low band is an element in one and only one vector or object. Various segmentation algorithms can be applied to the present invention, such as disclosed by Patrick Campbell McLean, in "Structured Video Coding", master thesis submitted to the Media Arts and Science Section, Massachusetts Institute of Technology, June 1991 or pending U.S. patent application Ser. No. 08/339,491 filed on Nov. 14, 1994.

Referring to FIGS. 14-15, the vector 1402 or object 1502 is propagated into higher bands by upsampling to define an

11

outline 1404 (1504) at the next finer scale. In order to maintain the vectors and/or objects at the same size as in the coarser scale, the coefficients of each vector or object within the outline are regrouped into four vectors 1406 or objects 1506 as illustrated in FIGS. 14–15. The vector or object at the coarser scale is called the vector or object parent and the four vectors or objects at the next finer scale are called vector or object children. The four vector/object children are extracted from the outline by taking the four polyphase components 1408 (1508). This is equivalent to a shifted subsampling of the outline. Examples of the vector parent-child relationship and object parent-child relationship are shown in FIGS. 14 and 15, respectively. These relationships define vector or object wavelet trees.

Vector or object wavelet trees are similar to the wavelet trees of scalar ZTE coding except that at each node of these vector or object wavelet trees there is a vector of wavelet coefficients instead of a single value. As is done in scalar ZTE coding described above, each tree is traversed depth-first from the root in the low-low band through the children and grandchildren. As nodes are traversed, coefficients are quantized according to the current quantization specification. This quantization can be done as a vector of scalar quantizations by scalar quantizer 1210 or as a single vector quantization by vector quantizer 1110. Various vector quantization techniques can be applied as disclosed by Allen Gersho et al. in "Vector Quantization and Signal Compression", Kluwer Academic Publishers, (1992).

Similar to the scalar ZTE coding as described above, it is beneficial to reorganize the coefficient vectors of each wavelet tree to form a "wavelet block" as shown in FIG. 16. To facilitate the depth-first scanning pattern, the invention may optionally reorganize the quantized coefficients of each wavelet tree to form a "wavelet block" 1604. As shown in FIGS. 11 and 12, the reorganization is accomplished in the vectors/objects wavelet tree reorganizer 1108 (1208) prior to quantization. Since reorganization is optional, both FIGS. 11 and 12 depict the optional nature of the reorganizer as path 1106 and path 1206 which respectively bypass the reorganizer and its associated function.

FIG. 16 schematically depicts a wavelet block 1604 that is generated by the invention. Each wavelet block 1604 of a frame comprises those coefficient vectors at all scales and orientations that represent the frame at the spatial location of that block organized in ascending frequency order from parent to children to grandchildren. Namely, the invention maps a tree 1602 extending from a vector 1600 in the low-low band 1606 (LL_3) in the wavelet tree 1602 into a wavelet block 1604. Such a structure permits the quantization factor, for the case of scalar quantization, or the codebook, for the case of vector quantization, to adapt to each block according to where the block is spatially located and thus what it represents in the frame.

By using a depth-first scanning pattern, each wavelet block is completely scanned to quantize its coefficient vectors before the next block is scanned and so on. For example, block 1610 is completely scanned, then block 1612, then block 1614, and so on in a raster scan pattern through the frame of wavelet blocks. The ordering of blocks does not have to be in a raster scan pattern, but can be any order as desired by the application.

Furthermore, this invention presents two distinct methods for performing the quantization of the vectors of wavelet coefficients. Quantization can be performed as scalar quantization on each element of each vector, or it can be performed as vector quantization on each vector. At each

12

node of each tree, a symbol assignor 1112 (1212) assigns a symbol to characterize the quantized vector at that node. The symbols and the quantized coefficient values are all encoded using an entropy coder 1114 (1214), such as an arithmetic coder.

After either type of quantization, zerotrees exist wherever the coefficient vector at a node is the zero vector and all descendants are zerotrees. The wavelet trees are efficiently represented and coded by assigning symbols to represent the nodes and by pruning the trees at zerotrees. An arithmetic coder is used to assign the minimum number of bits to only those symbols and values needed by the decoder to faithfully reconstruct the trees of wavelet coefficient vectors.

One embodiment of this invention (as shown in FIG. 12) uses scalar quantization of the elements of the vectors of wavelet coefficients. After quantization, each vector will contain either all zeros or one or more nonzero values. A vector zerotree exists where there is a zero vector at every node of the tree. A vector zerotree root is the root of a vector zerotree. The coefficient vector at a vector zerotree root is the zero vector and all descendants of the vector zerotree root are themselves vector zerotree roots.

Zerotree roots are important because wavelet trees can be pruned at zerotree roots and therefore no bits are needed to code the coefficients that have been pruned away. Therefore, to increase coding efficiency, it is desirable to have as many zerotrees after quantization as possible. One method is to employ a "post-quantization processor" 1211 (1011) to perform some post-quantization processing (as illustrated in FIGS. 10 and 12) of each vector in an effort to change it to the zero vector when doing so would cause only negligible additional distortion to the final reconstructed frame. A measure of each vector can be computed and when below some threshold, the vector would be set to zero.

After scalar quantization of the vector elements and post-quantization processing to force vectors to the zero vectors where appropriate, scanning and classification of the tree nodes is done in a manner similar to that in scalar ZTE coding. Each node is assigned one of three symbols: "zerotree root", "valued zerotree root", and "value". These symbols have the same meaning as in scalar ZTE coding, except that the nodes are vectors and the classification of zero or nonzero refers to the whole vector. An entropy coder is used to code the symbols and the tables of values. The tables of values are constructed as in scalar ZTE coding. To entropy code the values requires using vector entropy coding instead of scalar entropy coding.

One embodiment of this invention (as shown in FIG. 11) uses vector quantization of the vectors of wavelet coefficients. Vector quantization (VQ) is a mature field and various forms of VQ desired can be used. As discussed above, it is desirable to increase the number of zerotree roots when doing so does not significantly increase the distortion of the reconstructed frame. To achieve this result, the VQ can be adapted to incorporate a form of dead zone whereby vectors with a measure below some threshold are mapped to the index that reconstructs as the zero vector.

Vector quantization replaces the vector at each node of the tree with an index into a codebook of vectors that selects the best reconstruction vector for that node. After vector quantization, scanning and classification of the tree nodes is done in a manner similar to that in scalar ZTE coding. Each node is assigned one of three symbols: "zerotree root", "valued zerotree root", and "value". These symbols have the same meaning as in scalar ZTE coding, except that the nodes are now indices into a codebook and zero or nonzero refers

13

to the codebook entry for each index. An entropy coder is used to code the symbols and the tables of indices. The tables of indices are constructed as in scalar ZTE coding.

FIGS. 9 and 10 illustrate alternative embodiments of the present invention that use the vector wavelet transform. Again, since the encoders 900 and 1000 of FIGS. 9 and 10 are similar to the encoders of FIGS. 11 and 12 respectively, devices 908-914 and 1008-1014 are not discussed. Since these devices perform the same functions as those of 1108-1114 and 1208-1214 respectively, the descriptions for these devices are identical to the discussion above.

However, unlike encoders 1100 and 1200, the encoders 900 and 1000 employ a vector organizer 903 (1003) in conjunction with a vector wavelet tree generator 904 (1004), as discussed above with reference to FIG. 13. After the vector wavelet trees are generated, the encoders 900 and 1000 behave similarly to the encoders 1100 and 1200 respectively.

Those skilled in the art will realize that although the vector or object wavelet coding incorporates the Zerotree Entropy (ZTE) coding method, it can also be implemented with other coding method such as the Embedded Zerotree Wavelet (EZW) algorithm.

Although various embodiments which incorporate the teachings of the present invention have been shown and described in detail herein, those skilled in the art can readily devise many other varied embodiments that still incorporate these teachings.

What is claimed is:

1. Method for organizing a vector parent-child relationship, said method comprising the steps of:

- (a) propagating a vector parent into a higher band by upsampling to define an outline;
- (b) regrouping coefficients within said outline into a plurality of vector children; and
- (c) extracting said vector children by taking a shifted subsampling of said outline.

2. Method for organizing a vector parent-child relationship, said method comprising the steps of:

- (a) propagating a vector parent into a higher band by upsampling to define an outline;
- (b) regrouping coefficients within said outline into a plurality of vector children; and
- (c) extracting said vector children by taking a plurality of polyphase components of said outline by spatially offsetting and then subsampling in each of two dimensions.

3. The method of claim 1, wherein said regrouping step (b) regroups coefficients within said outline into four vector children.

4. The method of claim 1, wherein said propagating step (a) propagates an object parent into a higher band by upsampling to define an outline, and wherein said regrouping step (b) regroups coefficients within said outline into a plurality of object children.

5. Method for organizing a vector wavelet for a frame, said method comprising the steps of:

14

(a) decomposing the frame into a plurality of polyphase components, wherein said decomposing step decomposes the frame into a plurality of polyphase components by spatially offsetting the frame and then subsampling the offset frame in each of two dimensions;

(b) applying a transform to each of said plurality of polyphase components; and

(c) forming a vector wavelet by grouping together transform coefficients from each of said transformed polyphase components.

6. The method of claim 5, wherein said applying step (b) applies a wavelet transform to each of said plurality of polyphase components.

7. The method of claim 5, wherein said forming step (c) forms an object wavelet by grouping together transform coefficients from each of said transformed polyphase components.

8. Apparatus for organizing a vector wavelet for a frame, said apparatus comprising:

a decomposing means for decomposing the frame into a plurality of polyphase components, wherein said decomposing means decomposes the frame into a plurality of polyphase components by spatially offsetting the frame and then subsampling the offset frame in each of two dimensions;

a transforming means for transforming each of said plurality of polyphase components; and

a forming means for forming a vector wavelet by grouping together transform coefficients from each of said transformed polyphase components.

9. The apparatus of claim 8, wherein said transforming means is a wavelet transform.

10. The apparatus of claim 8, wherein said forming means forms an object wavelet by grouping together transform coefficients from each of said transformed polyphase components.

11. A computer-readable medium having stored thereon a plurality of instructions, the plurality of instructions including instructions which, when executed by a processor, cause the processor to perform the steps comprising of:

(a) decomposing a frame into a plurality of polyphase components, wherein said decomposing step decomposes the frame into a plurality of polyphase components by spatially offsetting the frame and then subsampling the offset frame in each of two dimensions;

(b) applying a transform to each of said plurality of polyphase components; and

(c) forming a vector wavelet by grouping together transform coefficients from each of said transformed polyphase components.

12. The computer-readable medium of claim 11, wherein said applying step (b) applies a wavelet transform to each of said plurality of polyphase components.

13. The computer-readable medium of claim 11, wherein said forming step (c) forms an object wavelet by grouping together transform coefficients from each of said transformed polyphase components.

* * * * *